

Original communication

Interpreting the color effect of melanin on cocaine and benzoylecgonine assays for hair analysis: Brown and black samples compared

Tom Mieczkowski^{a,*}, Michael Kruger^b

^a Department of Criminology, University of South Florida, 4202 East Fowler Avenue, SOC 107, Tampa, FL 33620, United States

^b C.S. Mott Center for Human Growth and Development, Department of Obstetrics and Gynecology, School of Medicine, Wayne State University, United States

Received 23 June 2005; received in revised form 15 September 2005; accepted 28 September 2005

Available online 13 October 2006

Abstract

This paper examines the hypothesis that cocaine and BE assays of hair demonstrate a putative bias effect for darker color hair samples. Although such an effect has been reported in the literature, no one has examined this claim outside the bounds of simple significance of mean difference. In this paper a large number of black and brown hair samples are compared for cocaine and BE concentration values, and this comparison is evaluated for both significance and effect size. Two innovative measures are used to assess this relationship – a calculation of effect size using Cohen's *d*, and the use of an ROC curve to evaluate the potential for a dark color bias. The paper reports mixed results for significance, but consistent results for effect size. There does not appear to be any significant effect for cocaine. While BE demonstrates a significant mean difference, both the effect size and the ROC analysis show the effect to be trivial.

© 2005 Elsevier Ltd and AFP. All rights reserved.

Keywords: Hair analysis; Hair color effect on assays; Hair analysis of cocaine; Hair color bias

1. Introduction

Over the last several years the issue of the importance of melanin concentration on the interpretation of hair assays has been the subject of several papers.^{1–3} This topic has stirred more than the usual concern since it has emerged as a legal strategy in litigation involving actions based on the results of drug tests. For example, individuals disciplined in the work place for a cocaine positive hair test have litigated that action and raised a defense that (among other issues) hair analysis is “biased” against persons of color or persons with dark hair.

The data offered in support of the position that a melanin “bias” has been demonstrated is very limited. While it is not controversial that melanin is a material which binds many drugs, including cocaine and its metabolite benzoylecgonine (BE), it is also well established that cocaine binds to a large number of other entities in hair, including keratin protein.⁴ It has been shown, for example, that cocaine can be readily recovered from hair which is lacking in any pigment, and it has been readily shown that cocaine can be found in the non-melanin fraction of hair after melanin has been removed by centrifugation.⁵

The basis of the argument regarding the potential for a melanin based “bias” would take on more importance if it were shown that the effect of melanin were routinely consequential for the outcome of an assay. That is, an assay which would otherwise be “negative” at some pre-determined threshold becomes “positive” due to the presence of melanin. Whether or not this would prove to be problematic in the use of hair assays is itself questionable. Even if it were shown to be so, thresholds could be adjusted for hair coloration much in the same way other kinds of assessments are contextually modified or interpreted.³ As well, sample preparation protocols may be employed that

* Corresponding author. Tel.: +1 813 974 8281; fax: +1 813 974 2803.
E-mail address: mieczkow@cas.usf.edu (T. Mieczkowski).

remove melanin by centrifugation prior to subjecting the sample to analysis.

Several studies have failed to find the putative “melanin bias” when examining large numbers of assay outcomes.^{6–8} In 2000 Mieczkowski and Newel¹ reported on a large number of data sets bearing on this point. They reviewed eight different data sets for the purposes of assessing the possibility that reported color of hair can produce a systematic bias in the interpretation of hair assays, including sets on heroin and its metabolites, cocaine and its metabolites, MDMA and its analogs, and amphetamine and methamphetamine. These studies utilized a variety of different degrees of color categorization, ranging from the simple dichotomy of brown and black, to a high of 12 categories. The mean number of categories reported was 6 (mean = 5.875) and the analysis was based on 2791 data points. They utilized two major statistical techniques for assessing significance; one-way analysis of variance, and Tukey’s Honestly Significant Difference procedure. The analysis of this data failed to discern a significant color effect.

One possible reason is that melanin forms only a small fraction of the materials which constitute hair. Another may be that concentration differences can only be detected when comparing extreme contrasts (high melanin hair vs. melanin-free hair). Another may be that the failure to remove melanin from the sample prior to analysis can heighten the concentration recovered. Ultimately, while melanin may cause some degree of differentiation, this effect may prove to be relatively trivial in most applications. Therefore, it is the magnitude of the *effect size* attributable to melanin that is the critical question at hand. Even if melanin binds cocaine, does it play a large enough role to cause a consequence in the outcome of the assays of persons who use cocaine? Does it lead one to conclude that an assay would likely be misinterpreted because the hair color of the person was not included in the assessment of the assay results?

2. Effect size

Studies to date have presented data, and in some cases calculated the significance value, of a contrasting hair color/cocaine assay concentration value by comparing the means of the distinct color categories. These reports have either employed an *f* statistic, using an analysis of variance approach or a *t* test for comparing means of independent samples. They have in some cases shown a significant difference in the mean concentration values of cocaine recovered when controlling for hair color. There have been criticisms of these studies’ statistical methodologies, primarily that the number of samples compared is so small as to be uninformative and inconsistent with the level of generalization postulated by the researchers. There have also been studies which have not been able to replicate these findings, notably when one tries to identify these effects in much larger sample aggregates. Most impor-

tantly, however, is that there has been little or no effort in the cases of findings of significant difference by hair color to assess the strength of the relationship.

Statistical significance at a particular *p* value does not measure the impact or “effect” that one variable has upon another. It measures the likelihood that the relationship is a chance occurrence, and allows one to interpret the precision of the relationship. But if there is a statistically significant relationship between melanin and cocaine assay values, the significance statistic does not allow one to answer the question “is this relationship a strong, moderate, or weak one”? One way to evaluate this question is to calculate an effect size.

The effect size is a measure of the consequences of an independent variable on a dependent variable. It is a measure of the magnitude of the predictor variable on a resultant variable and quantifies the impact of the predictor variable. In the case of evaluating differences in cocaine concentration between hair color groups, for example, it allows one to assign a value to the amount of explanatory power of the independent variable (hair color) on the dependent variable (analyte concentration). In this regard, it is a supplement to statistical significance, and should be calculated any time a statistically significant difference between means is reported.

The effect size is calculated as a standardized measure of the difference between means. The pooled SD is typically used as the standardization measure. There are several different approaches to calculating effect size. A widely employed measure is Cohen’s *d*. It is calculated as:

$$d = \frac{M_1 - M_2}{\sigma_{\text{pooled}}},$$

where $\sigma_{\text{pooled}} = \sqrt{[\sigma_1^2 + \sigma_2^2]/2}$ and *M* = group mean.

Cohen’s *d* is also related to the *t* statistic, used to compare two independent groups. It is also related to the point-biserial Pearson’s *r*. These measures can all be derived from each other. For example, it can be shown that $d = 2t/\sqrt{df}$ where *df* is the number of degrees of freedom for *t*. As well, $d = 2r/\sqrt{1 - r^2}$. The term *r* is used here to symbolize the point-biserial correlation between a dichotomous independent variable and a continuous dependent variable, a special case of the Pearson product-moment correlation. This special case of *r* is defined as $r_{y,z} = r_{dv,iv}$. There are a number of methods to arrive at the point-biserial correlation. Like *d*, *r* can be derived from *t* as it can be shown that $r = \sqrt{[t^2/(t^2 + df)]}$. Also, *r* can be directly calculated from *d* as $r = d/\sqrt{(d^2 + 4)}$.

2.1. The effect of color on cocaine and BE assay values

This analysis seeks to evaluate the degree to which both cocaine and BE assays of hair samples have differentiated outcomes attributable to color. A number of studies have attempted to document this effect, with different studies reporting different outcomes and interpretations. In this analysis, we focus on differentiating the outcomes for

brown vs. black hair. For each of these two color categories, we examine both the concentration values of GC/MS analyses of hair samples (expressed as ng/10 mg of hair sample) for cocaine and BE. We seek to answer the following questions:

1. Is there a significant difference between black and brown hair samples?
2. If there is a significant difference, how large is the effect size?
3. If such an effect is found, does it have an enhanced impact at the threshold value for a positive assay?
4. What interpretive implications do these analyses have for utilizing hair assays?

While many studies have contrasted hair colors of extreme difference (e.g., blond or white vs. black) this study compares black and brown hair samples. This comparison is of particular interest because these two color categories constitute approximately 90% of the natural hair color of persons in the United States⁹ and represent the likeliest confounder of interpretation in detecting color bias at the threshold. Indeed, several papers claim to have identified a significant or important distinction between black and brown hair samples and cocaine concentration.^{10,11} These findings have been cited in supporting the “color bias” hypothesis.

About 56.1% of Americans are estimated to have natural brown hair and 33.6% to have natural black hair.^A For these two color groups we present a frequency distribution, and a *t* test comparing the group means for the cocaine concentrations. If the means have a significant difference at $p \leq 0.05$, we calculate an effect size value, Cohen's *d*. Additionally, we utilize receiver operating characteristic (ROC) analysis to determine if the use of one color as a state variable can discern positive and negative assay outcomes.

In addition to examining the overall outcome patterns, we also present a focused analysis of the distribution of black and brown hair samples at the region of the cutoff threshold of 5 ng/10 mg of hair. Some researchers have suggested that the problems of bias are likely to be most evident at the region of the threshold. Cases of extreme use of cocaine, for example, are recognized as incontrovertibly identifiable since the assay values will be extremely high. If the alleged “bias” regarding black vs. other hair colors is a factor, it is argued to be critical in the region where the cutoff is established. It (black hair pigmentation) may contribute enough to be a “tipping factor” pushing these samples above the threshold while non-black samples remain slightly below the threshold. The study examines the pattern of black and brown samples in two narrow ranges (± 2 and ± 1 ng) around the threshold value of 5 ng/10 mg.

3. Data and methods

Data for this study come from the Psychomedics Corporation, a large commercial hair analysis laboratory located in Los Angeles, CA. The data set consist of 8687 hair samples routinely received by the laboratory for analysis from corporate clients around the United States. These samples all had cocaine and BE positive assay at the limit of detection of (2 ng/10 mg) or greater. The samples were collected over the course of approximately 1 year. The data set consists solely of the assay values and the color category of the hair. The Psychomedics analysis protocol has been published elsewhere, and we will not describe it in detail here.¹² In brief, samples were weighed, digested and assayed using radioimmunoassay of hair digests. All samples were subject to a washing procedure as follows: First, dry isopropanol (2 mL) was added to about 12 mg of hair in 12 mm \times 75 mm tubes; the tubes were shaken vigorously at 37 °C for 15 min; then the isopropanol was removed to a separate tube and saved for later analysis. Then 2 mL of 0.01 M phosphate buffer/0.01% BSA, pH 6, was added to the hair samples in the tubes and the tubes shaken vigorously for 30 min at 37 °C, after which the buffer was removed and saved to a another tube for later analysis. This 30 min wash was repeated twice more, followed by two 60 min washes using the same conditions. After the final (5th) phosphate buffer wash and removal of the buffer, the hair sample was enzymatically digested prior to RIA analysis or confirmation by GC/MS or LC/MS/MS.

For cocaine and BE confirmation analyses, washed hair was enzymatically digested for 6 h at pH 5.5 for cocaine and metabolites and pH 6.65 for opiates. Samples for cocaine and BE analysis were extracted using Isolute SPE columns and derivatized to enhance chromatographic separation of BE from non-cocaine. This was accomplished by converting BE into a *n*-propyl derivative, followed by the LC/MS/MS analysis. A triple quadrupole API 2000 Perkin–Elmer Sciex (Thornhill, Ont., Canada) MS equipped with an atmospheric pressure ionization source via an ion-spray interface was used in all measurements. For LC, a binary pump with a Perkin–Elmer series 200 autosampler was used. The high performance (HPLC) column was a Keystone Scientific BETASIL C8, and the HPLC mobile phase consisted of water and acetonitrile containing 0.1% formic acid. The proportion of water/acetonitrile was 80:20. The MS was optimized for cocaine metabolites to give optimum ion yields. Ionization of analytes was obtained in the positive mode. Positive ion precursors for cocaine (*m/z* 304), BE (*m/z* 332), CE (*m/z* 318) and NOR (*m/z* 290) and their deuterated internal standards cocaine D3 (*m/z* 307), BE-D3 (*m/z* 335), CE-D3 (*m/z* 321) and NOR-D3 (*m/z* 293) were selected as the target analytes through the first quadrupole (Q1). Nitrogen was used as the collision gas in the second quadrupole (Q2). The product ions monitored in the third quadrupole (Q3) were *m/z* 182 and *m/z* 185, for cocaine and cocaine-D3 respectively; *m/z* 210 and *m/z* 213 for BE and BE-D3, respectively; *m/z*

^A Slightly less than 2% are estimated to have red hair, and slightly more than 8% are estimated to be naturally blond.

Table 1
Cocaine concentration of black and brown samples, raw data and log-transformed data

Hair color	Black (<i>N</i> = 4911)	Brown (<i>N</i> = 3367)
<i>Raw data</i>		
Mean cocaine concentration (ng/10 mg)	25.53	23.98
SD	38.66	36.26
Median	7.9	8.4
<i>Log-transformed data</i>		
Mean cocaine concentration, log-transformed (ng/10 mg)	0.9800	0.9759
SD	0.63096	0.61075
Median	0.8970	0.9243

196 and *m/z* 199 for CE and CE-D3, respectively, and *m/z* 168 and *m/z* 171 for NOR and NOR-D3, respectively. The values reported here are based on the GC/MS confirmatory tests. The concentration values are in ngs of cocaine or BE/10 mg of sample. The threshold conventionally applied to differentiate a positive and negative outcome is 5 ng/10 mg of sample.

3.1. Data analysis: contrasting black and brown hair samples

3.1.1. The distribution of cocaine

The first series of contrasts compares the total number of black and brown hair samples for cocaine concentration. Table 1 reports the number of samples in each category, their mean concentration values, SDs, and median values. Because the data do not constitute a normal distribution,^B we also do a log transformation and report the log-transformed values. In cases of drug testing, the nature of the outcome from employment field data is virtually always negatively skewed. The log transformation expresses the values at the base-10 logarithm of the concentration values.

The log transformed concentration values more closely approximate a normalized distribution. Fig. 1 displays the histogram of all raw data cocaine assay values for both black and brown hair color and shows a decidedly negative skew towards lower concentration values. Fig. 2 shows the histogram of the log-transformed values.

The results of a *t* test comparison of the mean cocaine concentration values for all black and brown hair samples – for both raw and the log-transformed data – shows no significance with $t_{\text{obt}} = 1.836$ with $p = 0.066$ for the raw data and $t_{\text{obt}} = 0.295$ with $p = 0.768$ for the log-transformed data. Since the relationship is not significant, no effect size is calculated. Figs. 3 and 4 display the log-transformed histograms of black hair samples (Fig. 3) and brown hair samples (Fig. 4) with the dashed line indicating the log of the threshold value and each graph showing a superimposed normal curve.

^B One sample K-S (raw data) = 23.35, $p = 0.000$; K-S (log-transformed data) = 4.661, $p = 0.000$.

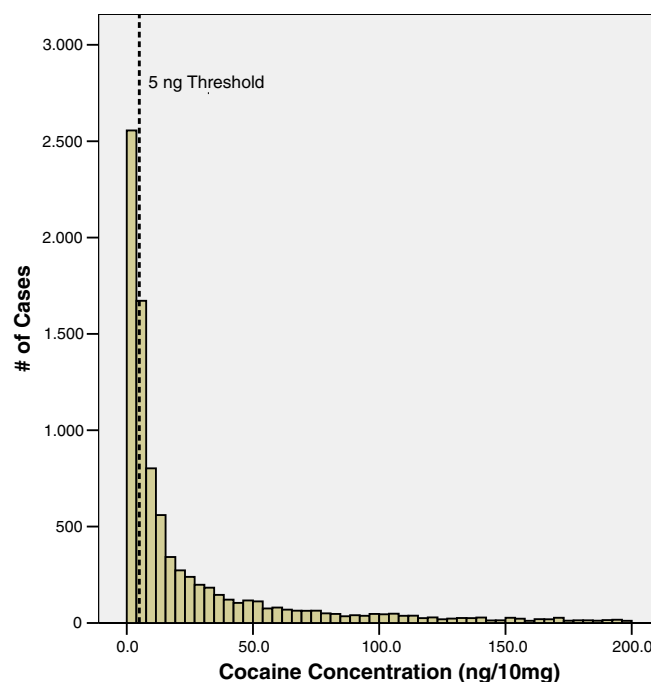


Fig. 1. Histogram of cocaine assay values, all samples.

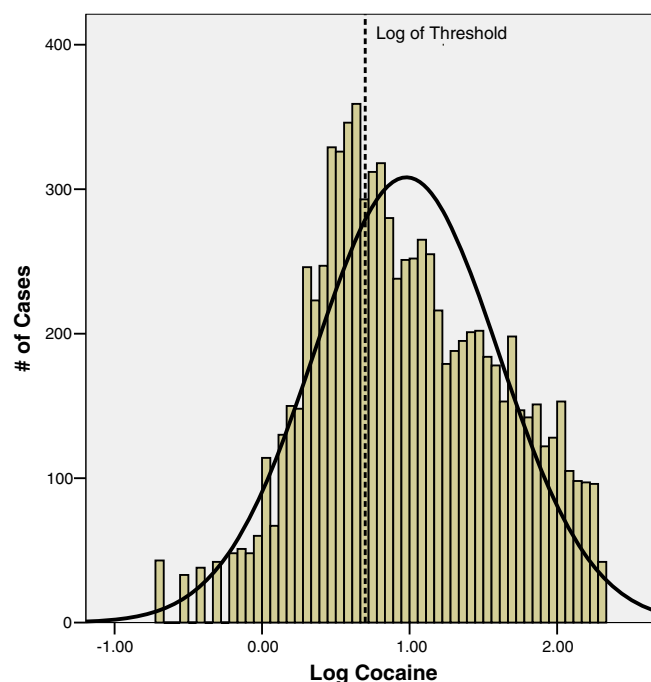


Fig. 2. Cocaine assays, all samples, log-transformed.

3.1.2. The distribution of cocaine at the threshold value

As previously noted, it has been suggested that the issue of hair color as a potential biasing effect is of special concern at the threshold or cutoff value of assays used for interpretative purposes. That is, a variance by color being systematically present would produce a discernible effect resulting in more black hair samples “crossing” the threshold in a narrow critical range at the minimum concentration

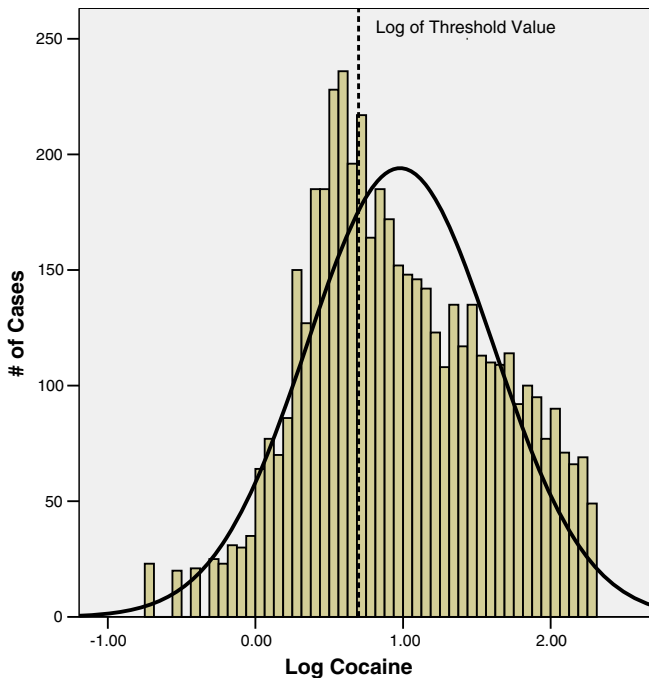


Fig. 3. Log-transformed cocaine concentrations, black samples.

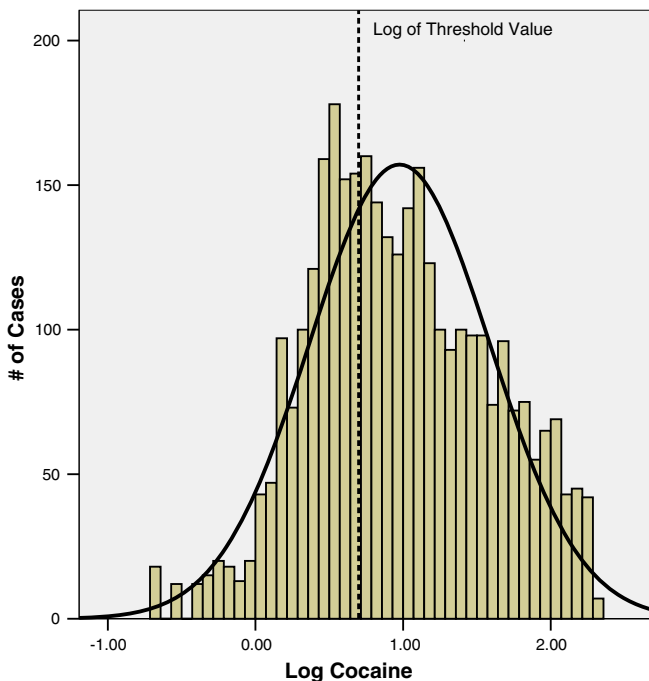


Fig. 4. Log-transformed cocaine concentrations, brown samples.

value, hence possibly producing an evidentiary positive sample. It is of little concern at extremely high or low values, which would not be “overwhelmed” by a color effect. In this study we have considered 5 ng/10 mg of sample to define that threshold. Table 2 compares the cocaine assay values considered in the range of 3–7 ng/10 mg of hair (± 2 U clustered around the cutoff of 5 ng/10 mg) shows no significant difference between black and brown hair.

Table 2

Cocaine concentration of black and brown samples in the 3–7 ng/10 mg range, raw data and log-transformed data

Hair color	Black (<i>N</i> = 1237)	Brown (<i>N</i> = 847)
<i>Raw data</i>		
Mean cocaine concentration (ng/10 mg), threshold ± 2	4.622	4.659
SD	1.1380	1.1837
Median	4.500	4.400
<i>Log-transformed data</i>		
Mean cocaine concentration, log-transformed (ng/10 mg), threshold ± 2	0.6518	0.6544
SD	0.10621	0.11015
Median	0.6532	0.6435

For the raw data $t_{\text{obt}} = -0.726$, $p = 0.596$ and for the log-transformed data $t_{\text{obt}} = -0.534$, $p = 0.593$.

Likewise if the range is narrowed further, to ± 1 ng/10 mg of sample, the relationship remains non-significant. For the raw data $t_{\text{obt}} = -0.966$, $p = 0.334$ and for the log-transformed data $t_{\text{obt}} = -0.906$, $p = 0.365$. Table 3 reports the values for each color group clustered at the 5 ng/10 mg cut-off value in the ranges ± 1 ng.

Fig. 5 displays the percentage frequency distribution of these cocaine assays for the ± 1 ng range. Fig. 5 standardizes the cases as percentages, since the sample sizes are not equal. The tracking of the two groups in the range are essentially parallel. This pattern is not changed by expanding the range to a ± 2 ng interval, so the graph is not shown here.

Since there are no statistically significant differences for either of these comparisons, no effect size is calculated.

3.1.3. Benzoylcegonine comparisons

3.1.3.1. The distribution of BE. Table 4 reports the number of samples in each category, their mean BE concentration values and SDs.

A t test comparison of all black and brown sample mean values for BE shows that this difference is significant ($t_{\text{obt}} = 2.883$, $p = 0.004$). Fig. 6 displays the raw data frequency distribution by percentage of the BE concentration values for black and brown samples. The vertical line

Table 3

Cocaine concentration of black and brown samples in the 4–6 ng/10 mg range, raw data and log-transformed data

Hair color	Black (<i>N</i> = 619)	Brown (<i>N</i> = 394)
<i>Raw data</i>		
Mean cocaine concentration (ng/10 mg), threshold ± 1	4.890	4.928
SD	0.5963	0.6197
Median	4.80	4.90
<i>Log-transformed data</i>		
Mean cocaine concentration, log-transformed (ng/10 mg), threshold ± 1	0.6861	0.6892
SD	0.05282	0.05457
Median	0.6812	0.6902

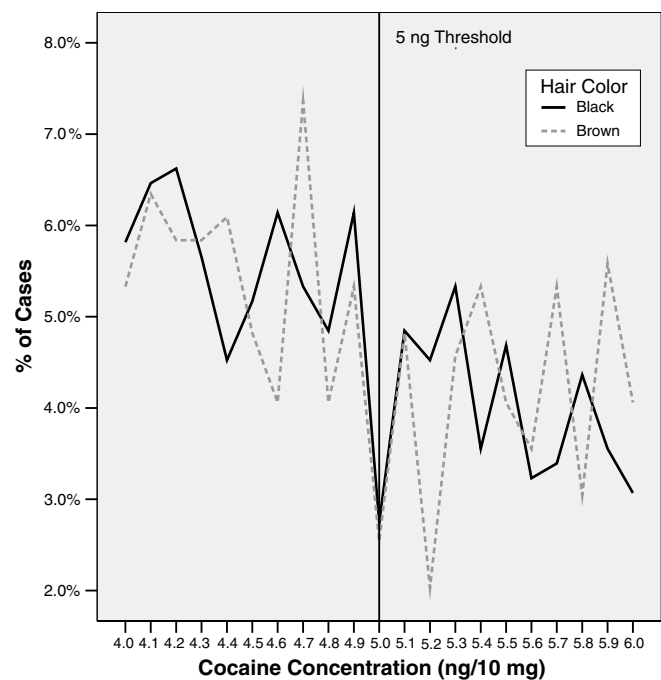


Fig. 5. Cocaine values at ± 1 ng of threshold.

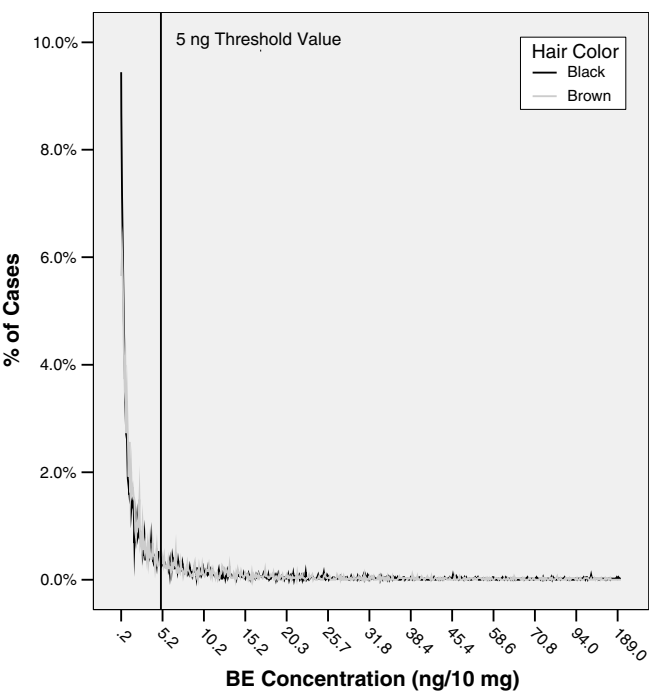


Fig. 6. BE concentration, black and brown samples.

Table 4
BE concentration of black and brown samples

Hair color	Black (<i>N</i> = 4246)	Brown (<i>N</i> = 3079)
<i>Raw data</i>		
Mean BE concentration (ng/10 mg)	8.104	6.896
SD	19.47	14.93
Median	1.50	1.70
<i>Log-transformed data</i>		
Mean BE concentration, log-transformed (ng/10 mg)	0.2861	0.3122
SD	0.7114	0.6519
Median	0.1761	0.2304

marks the threshold value of 5 ng/10 mg. Fig. 7 displays the log-transformed values for BE.

Since we have a significant effect for the difference between the black and brown hair for BE, we can calculate an effect size for the BE/color relationship. We calculate Cohen's $d = 0.0696$, with an associated Pearson's biserial r equal to 0.034 (significant at $p = 0.01$). The measures of effect size are summarized in Table 5. In addition to Cohen's d the table displays two interpretative measures of the effect size. One is the percentage of variance in the dependent variable (BE concentration) accounted for by the independent variable (hair color). The second is to conceptualize Cohen's d as an index of the degree of overlap between the two distributions. Table 5 indicates there is virtually complete overlap between the two distributions.

Consideration of the degree of overlap is an alternative way to interpret Cohen's d . A d of 0.0 would indicate perfect coincidence or overlap of the two distributions (i.e., they have an overlap of 100%). Cohen's d increases in value

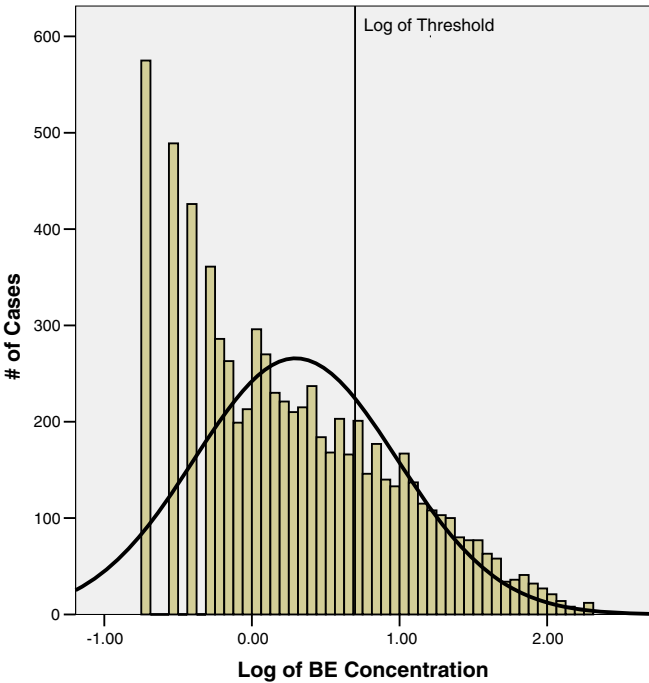


Fig. 7. Log-transformed BE concentration, black and brown samples.

Table 5
Measures of the effect size for BE and hair color

Measure	Outcome
Cohen's d	0.0696
% Explained variance by hair color	0.12%
Distribution Overlap	Approaching 100%

as the distribution move apart and become distinct. At a d of 2.0, for example, only about 15% of the areas of the distributions overlap. The less overlap the more distinction between the comparison groups. Cohen's d which falls below 0.1 indicates practically total (or 100%) overlap and no distinction between comparison groups.

The interpretation of the effect size, as previously discussed, can be expressed in a number of ways. A common method is to express the effect as a coefficient of the variance explained by the independent variable (IV) in the dependent variable (DV); in essence, the amount of change in the DV corresponding to a unit change in the IV. In this case that changes amount to 0.0012. The practical consequences of this, for example, would lower the threshold of detection for BE by 0.006 ng/mg from 5.0 to 4.994. An analysis of cases at the threshold range (4.0–5.0) shows that this would produce no change in outcome status. The threshold value is reported as 4.9, so a value of 4.994 is measured as a 4.9, and consequently a negative. Because the explained variance is a magnitude lower than the reported values (which are reported to the nearest tenth), the effect on cutoff decisions is practically non-existent.

3.1.3.2. BE values at the 5 ng threshold. A comparison of the black vs. brown hair mean values for BE at the 5 ng threshold reveals no statistical significance by hair color at ranges of ± 2 ng or ± 1 ng. If BE values are considered in the range of 3–7 ng/10 mg of hair there is no significant difference between black and brown hair ($t_{\text{obt}} = 0.260$, $p = 0.795$). Likewise, if the range is narrowed further, to ± 1 ng/10 mg of sample, the relationship remains non-significant ($t_{\text{obt}} = 0.252$, $p = 0.801$). Table 6 reports the values for each color group clustered at the 5 ng/10 mg cutoff value in the ± 1 ng range.

Fig. 8 displays the BE distribution frequencies at 5 ng/10 mg of sample within the range of ± 1 ng/10 mg. The frequencies are expressed as percentages to standardize the unequal N's of the two groups.

The pattern in Fig. 8 repeats for BE the pattern observed for cocaine in that the concentration value trends are generally parallel. The percentage of brown samples is somewhat more peaked at the threshold than black samples, but these fluctuations are not statistically meaningful.

Table 6
Concentration of BE for black and brown samples at threshold ± 1 ng

Hair color	Black ($N = 302$)	Brown ($N = 236$)
<i>Raw data</i>		
Mean BE concentration (ng/10 mg), threshold ± 1	4.899	4.886
SD	0.5823	0.5974
Median	4.80	4.89
<i>Log-transformed data</i>		
Mean BE concentration, log-transformed (ng/10 mg), threshold ± 1	0.6971	0.6858
SD	0.05140	0.05320
Median	0.6812	0.6902

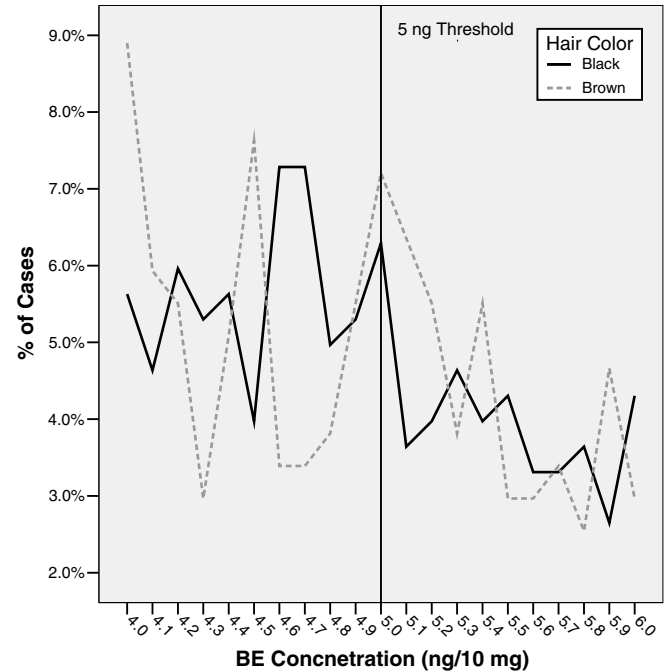


Fig. 8. BE concentrations at ± 1 ng of threshold.

Graphing the broader ranged interval of 3–7 ng/10 mg does not alter the relationship.

Table 7 reports the comparison of the mean BE concentration values in the 3–7 ng/10 mg range for black and brown samples. The differences are not significant. A comparison of means test produces a $t_{\text{obt}} = 0.260$ with $p = 0.795$.

3.1.4. ROC analysis of black color as a biasing factor

Another approach to evaluating the potential distinctions between black and brown hair samples is to utilize an ROC analysis. An ROC curve is a plot of the true positive rate against the false positive rate for the threshold value of a diagnostic test. An ROC analysis utilizes a continuous ratio or interval variable and allows for the assignment of a dichotomous nominal variable (the “state variable”) to one class. The assignment can either be “correct” (a true positive – defined as the variable’s sensitivity) or incorrect (a true negative – defined as the variable’s specificity). If we consider black hair color to be the state

Table 7
Concentration of BE for black and brown samples at threshold ± 2 ng

Hair color	Black ($N = 616$)	Brown ($N = 468$)
<i>Raw data</i>		
Mean BE (ng/10 mg), threshold ± 2	4.695	4.677
SD	1.1639	1.1603
Median	4.60	4.50
<i>Log-transformed data</i>		
Mean BE concentration, log-transformed (ng/10 mg), threshold ± 2	0.6583	0.6566
SD	0.1080	0.1078
Median	0.6628	0.6532

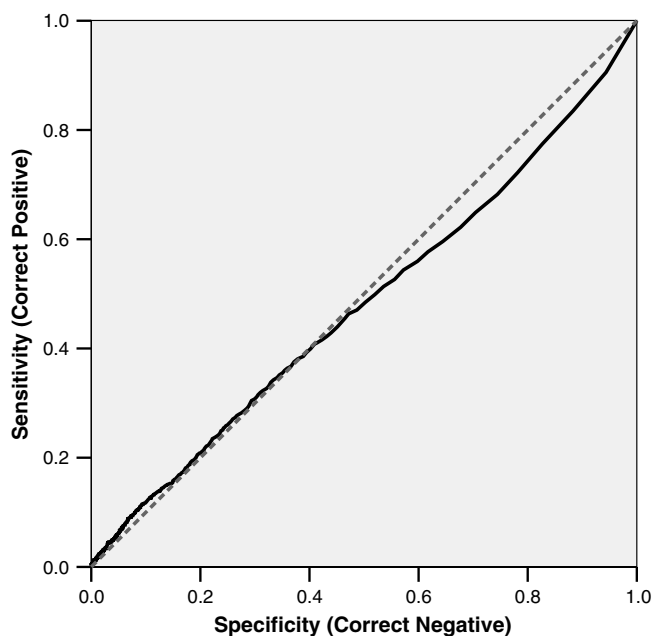


Fig. 9. The ROC curve for BE: black hair as the state variable.

variable and BE concentration to be the continuous variable, the ROC curve indicates to what degree the color state produces a predictable change between the value of the continuous variable and the class or category of the state variable. Here, this relationship is characterized as positive (black hair results in higher concentration values). Using a reduction of error approach, the informational contribution of the state variable should therefore result in an improvement over random chance assignment. The chance outcome for a dichotomous state variable is represented by a diagonal line from the origin and ascending upwards and to the right. For a dichotomous variable the probability value is 0.5. If the ROC describes an arc above this diagonal line, it is a measure of the effect on the state variable by the continuous variable. The larger the area between the arc and the diagonal reference line, the more pronounced the effect. The ROC curve for BE is shown in Fig. 9.

An examination of Fig. 9 shows that the ROC curve is nearly coincident on the diagonal, which indicates that the state variable (black hair color) as a predictor offers no better criteria for predicting BE concentration value than random chance.

4. Discussion

The issue of a “color bias” has received considerable attention in the literature on hair analysis. Generally, the findings which are consistent with this interpretation of the consequences of melanin concentration on cocaine retention in human and animal hair establish that melanin can bind cocaine and can play a role in influencing the concentration values for any particular assay approach. However, such studies employing some control mechanism by

dosage titration or similar approaches have largely relied on animal models. They have also contrasted pigmented and non-pigmented hair and not looked at a continuum of coloration, and have had very small sample numbers. There are few controlled dose studies of human subjects, and those which have been done examining cocaine concentration support the interpretation that “simple differences in melanin content are an unlikely explanation for the differences observed”.¹³ Kidwell,¹⁴ for example, discussing the issue of hair color effects stated that:

“We had termed this unequal uptake of drugs a matrix bias which grew into more popular usage as a racial bias. A number of studies have attempted to link this matrix bias with binding drugs to melanin. However, the results do not appear to be logically consistent.” (p. 10)

Studies which have examined very large data sets have confirmed this observation. There are a number of possible reasons for why a number of studies appear to produce conflicting findings. One of critical importance is that these studies have in the main used distinctly different sample preparation protocols, including such important preparatory steps as washing the samples of contaminants, removing the melanin fraction by centrifugation, and using distinctly different extraction methods.

In addition, these studies have utilized a variety of statistical procedures to determine significance, but have not reported an effect size when they have found a significant relationship between the drug concentration and hair color. Two problems are evident because of this; first, the small number of samples makes the data very unstable, and second the failure to report an effect size means that the potential bias cannot be assessed in interpreting the assay outcomes. In terms of data instability, a few cases or even a single case can completely reverse the statistical findings. These studies have consistently relied on mean values as comparators, which are sensitive to extreme scores. The failure to calculate an effect size means that although a relationship may be significant, one has no way to determine its actual impact in an interpretation of the assay in a clinical sense.

Regarding clinical or interpretive utility, the consistent pattern of contrasting pigmented and non-pigmented hair probably has little further to contribute to illuminating the effect of melanin. In real practice, few persons have non-pigmented hair. The overwhelming coloration of hair amongst humankind is black and brown. To date, there are no other researchers who have contrasted these two common color groups to see if any systematic differences exist between these groups. Yet these color groups constitute the overwhelming number of samples received by a practicing laboratory.

This paper reports a mixed outcome regarding significance. In reference to the first research question of the difference between black and brown hair samples the results are mixed. There is no significant difference for cocaine concentration by color group, but there is a significant difference for BE. However, in reference to the second

question on the effect of such a difference, we note that the effect size for BE is so small as to be trivial. In an attempt to determine whether the effect exerts some particular mechanics at the threshold (which constitutes the third question), the narrow range around the 5 ng/10 mg was examined. We found no significant differences between the two color groups within either the ± 1 or ± 2 range. An additional level of analysis, using an ROC “reduction of error” approach, shows there is no gain in predictive information by knowing the color, at least in terms of comparing black to brown samples.

In summary, regarding the question of a “color bias” for hair, these findings indicate that the issue needs to be rephrased. Instead of the simple notion of a “bias”, one should be asking “is there an importance or demonstrable consequence to hair color differentials when considering the use of hair analysis in various settings”? If there is, it seems quite possible to make some relatively simple clinical or assessment adjustments, such as modifying the threshold values to accommodate any small effect. It is possible that in cases of extreme lack of hair pigmentation, an adjustment may need to be made. Based on the data presented here, that need for most cases is not likely to be necessary.

Acknowledgment

The author would like to acknowledge the contributions and critical comments of Virginia Hill.

References

- Mieczkowski T, Newel R. A statistical examination of hair color as a potential biasing factor in hair analysis. *Forensic Sci Int* 2000;**107**:13–38.
- Hubbard D, Wilkins D, Rollins D. The incorporation of cocaine and metabolites into hair: effects of dose and hair pigmentation. *Drug Metab Dispos* 2000;**28**(12):1464–9.
- Borges C, Roberts J, Wilkins D, Rollins D. Cocaine, benzoylecgonine, amphetamine and *n*-acetylamphetamine binding to melanin subtypes. *J Anal Toxicol* 2003;**27**:125–34.
- Stout P, Ruth J. Deposition of ³Hcocaine, [³H]nicotine, and [³H]flunitrazepam in mouse melanosomes after systemic administration. *Drug Metabol Dispos* 1999;**27**(6):731–5.
- Baumgartner W, Hill V. Hair analysis for organic analytes: methodology, reliability issues, and field studies. In: Kintz P, editor. *Drug testing in hair*. New York: CRC Press; 1996. p. 223–66.
- Hoffman BH. Analysis of race effects on drug-test results. *J Occup Environ Med* 1999;**41**:612–4.
- Mieczkowski T, Kruger M. Analysis of color effects for hair analysis in a large sample: applications of general linear models. *Bull Int Assoc Forensic Toxicol* 2001;**31**(1):9–11.
- Kelly R, Mieczkowski T, Sweeney S, Bourne J. Hair analysis for drugs of abuse: hair color and race differentials or systematic differences in drug prevalence. *Forensic Sci Int* 2000;**107**:63–86.
- The Third National Health and Nutrition Examination Survey (NHANES III, 1988–94), National Center for Health Statistics, Centers for Disease Control (NCHS/CDC), Atlanta, GA, October, 1996.
- Kidwell D. What does hair analysis for drugs of abuse measure? *SAFS Newsletter* 1996;**XXIV**(1):9–18.
- Henderson G, Harkey M, Zhou C, Jones R, Jacob P. Incorporation of isotopically labeled cocaine into human hair: race as a factor. *J Anal Toxicol* 1998;**22**:156–65.
- Cairns T, Hill V, Schaffer M, Thistle W. Removing and identifying drug contamination in the analysis of human hair. *Forensic Sci Int* 2004;**145**:97–108.
- Henderson G, Harkey M, Zhou C, Jones R, Jacob P. Incorporation of isotopically labeled cocaine and metabolites into human hair: 1 dose–response relationships. *J Anal Toxicol* 1996;**20**:1–12.
- Kidwell D. Is hair testing culturally biased and why is this a concern in the U.S.? In Second European Meeting on Hair Analysis, Martigny, Switzerland, June, 1999.